

The Methodological Challenges of Measuring Student Learning, Degree Attainment, and Early Labor Market Outcomes in Higher Education

Tatiana Melguizo, Gema Zamarro, Tatiana Velasco, and Fabio J. Sanchez

Assessing the quality of higher education is hard but there is growing pressure for governments to create a ranking system for institutions that can be used for assessment and funding allocations. Such a system, however, would require a reliable methodology to fairly assess colleges using a wide variety of indicators. Countries with centralized governance structures have motivated researchers to develop “value-added” metrics of colleges’ contributions to student outcomes that can be used for summative assessment (Coates, 2009; Melguizo & Wainer, 2016; Shavelson et al. 2016). Estimating the “value-added” of colleges and programs, however, is methodologically challenging: first, high- and low-achieving students tend to self-select into different colleges— a behavior that if not accounted for, may yield to estimates that capture students’ prior achievement rather than colleges’ effectiveness at raising achievement; second, measures considering gains in student learning outcomes (SLOs) as indicators at the higher education level are scant. In our paper, we study these challenges and compare the methods used for obtaining value-added metrics in the context of higher education in Colombia.

How to best estimate value-added models in higher education?

We estimate ordinary least squares models that measure how much, on average, students in a given college and program perform above those in other colleges and programs that serve similar students. To address self-selection we control for the student’s score in the mandated standardized high school exit exam, the average score of their peers in the entering cohort, and other student characteristics. We compare the performance of the model when estimated using three different statistical methods: fixed effects (FE), random effects (RE) and aggregated residuals (AR). We estimate college-program value added on students’ scores in the generic component of a standardized college exit exam, graduation, employment in the formal sector, and initial wages. We find that FE methods are the most stable and robust to the issue of self-selection into programs and colleges.

Full Article Citation:

Melguizo, T., Zamarro, G., Velasco, T., & Sanchez, F. J. (2017). [The Methodological Challenges of Measuring Student Learning, Degree Attainment, and Early Labor Market Outcomes in Higher Education](#). *Journal of Research on Educational Effectiveness*, 10(2), 424–448. DOI: 10.1080/19345747.2016.1238985

Does student’s self-selection really matter? Yes, especially in SLOs. When we implement FE models to estimate gains in SLOs using the college-exit exam, we find that once we introduce controls for selection the variance in the value-added distribution is halved (see Figure below), indicating colleges’ contributions to the generic knowledge evaluated in the college-exit exam were mostly driven by selection of more prepared students into certain colleges and programs. However, longer term employment outcomes were less sensitive to the selection correction.

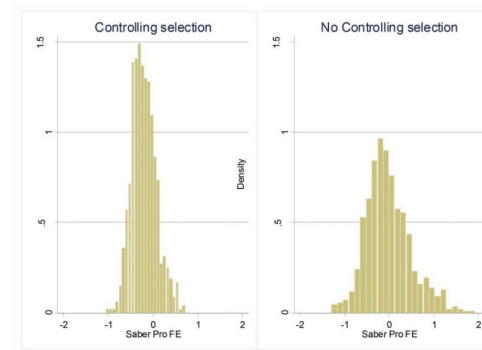


Figure. Distribution of fixed effects college-program contributions including and excluding selection controls

Are there differences in colleges value added across indicators?

Our findings indicate that rankings of specific college-program combinations change depending on the different outcomes considered. For example, we found that programs like math and natural sciences, added value both in terms of skills and employment, while others, such as in the agriculture and veterinary area, added little to students’ generic skills while still contributing to the probability of employment.

What did we learn? Any set of indicators developed should control for initial selection of students into colleges and programs. Researchers and policy makers need to continue to work to fine-tune these methodological tools before they can be used in any high-stakes summative evaluation.